

604006

Office of Naval Research

Contract Nonr-1866(16)

NR - 372 - 012

**A BAYESIAN APPROACH TO PROBLEMS IN  
STOCHASTIC ESTIMATION AND CONTROL**

by

**Yu-Chi Ho and Robert C. K. Lee \***

June 9, 1964

7p \$1.00 lc  
\$0.50 mf

The research reported in this document was supported by the U. S. Army Research Office, the U. S. Air Force Office of Scientific Research, and the U. S. Office of Naval Research under the Joint Services Electronics Program by Contract Nonr-1866(16), and by the Aeronautical Division of the Minneapolis-Honeywell Regulator Company, Boston, Massachusetts. Reproduction in whole or in part is permitted for any purpose of the United States Government.

\*Minneapolis-Honeywell Regulator Co., Boston, Massachusetts

**Technical Report No. 448**

**Cruft Laboratory**

**Division of Engineering and Applied Physics**

**Harvard University**

**Cambridge, Massachusetts**

# A BAYESIAN APPROACH TO PROBLEMS IN STOCHASTIC ESTIMATION AND CONTROL\*

Yu-Chi Ho\*\*  
Harvard University, Cambridge 38, Massachusetts

and

Robert C. K. Lee  
Minneapolis-Honeywell Regulator Company  
Boston, Massachusetts

## SUMMARY

In this paper, a general class of stochastic estimation and control problems is formulated from the Bayesian Decision-Theoretic viewpoint. A discussion as to how these problems can be solved step-by-step in principle and practice from this approach is presented. As a specific example, the closed form Wiener-Kalman solution for linear estimation in gaussian noise is derived. The purpose of the paper is to show that the Bayesian approach provides: (i) a general unifying framework within which to pursue further researches in stochastic estimation and control problems, (ii) the necessary computations and difficulties that must be overcome for these problems. An example of nonlinear, non-gaussian estimation problem is also solved.

## SINGLE STAGE ESTIMATION PROBLEM

For purpose of illustrating the concepts involved, the single stage estimation problem will be discussed first. Once this is accomplished, the multistage problem can be treated straightforwardly.

### Problem Statement

The following information is assumed given -

- (i) A set of measurements  $z_1, z_2, \dots, z_k$  which are denoted by the vector  $z$ .
- (ii) The physical relationship between the state of nature which is to be estimated and the measurements. This is given by

$$z = g(x, v) \quad (1)$$

where  $z$  is the measurement vector ( $k \times 1$ )

$x$  is the state (signal) vector ( $n \times 1$ )

$v$  is the noise vector ( $q \times 1$ )

- (iii) The joint density function  $p(x, v)$ :

From this one readily obtains the respective marginal density functions,  $p(x)$  and  $p(v)$ .

It is assumed that information for (iii) is available in analytical form or can be approximated by analytical distributions. Item (ii) can be either in closed form or merely computable. The problem is to obtain an estimate  $\hat{x}$  of  $x$  and which base upon the measurements is best in some sense to be defined later.

\*The work reported in this paper is supported in part by NONR Contract (1866)(16) at Harvard University and by Minneapolis-Honeywell Regulator Company, Aeronautical Division, Boston, Massachusetts.

\*\*Consultant at Minneapolis-Honeywell Regulator Company.

## The Bayesian Solution

The Bayesian solution to the above problem now proceeds via the following steps:

- (i) Evaluate  $p(z)$  - This can be done analytically, at least in principle, or experimentally by Monte Carlo methods since  $z = g(x, v)$  and  $p(x, v)$  are given. In the latter case, we assume it is possible to fit the experimental distribution again by a member of a family of distributions.
- (ii) At this point, two alternatives are possible, one may be superior to the other dependent on the nature of the problem.

- a) Evaluate  $p(x, z)$  - This is possible analytically if  $v$  is of the same dimension as  $z$  and one can obtain the functional relationship  $v = g^*(x, z)$  from (i) above. Then using  $p(x, v)$  and the theory of derived distributions, one obtains
 
$$p(x, z) = p(x, v = g^*(x, z)) |J| \quad (2)$$

where

$$J = \det \left[ \frac{\partial g^*(x, z)}{\partial z} \right]$$

- b) Evaluate  $p(z/x)$ . This conditional density function can always be obtained either analytically whenever possible or experimentally from the  $z = g(x, v)$  and  $p(x, v)$ .

Note that (iia) may be difficult to obtain in general since  $g^*$  may not exist either because of the non-linear nature of  $g$  or that  $z, v$  are of different dimensions. Nevertheless, (iib) can always be carried out. This fact will be demonstrated in the nonlinear example in the sequel.

- (iii) Evaluate  $p(x/z)$  using the following relationships,

- a) Following (iia)

$$p(x/z) = \frac{p(x, z)}{p(z)} \quad (3)$$

- b) Following (iib), use the Bayes' rule

$$p(x/z) = \frac{p(z/x) p(x)}{p(z)} \quad (4)$$

Depending on the class of distributions one has assumed or obtained for  $p(x, v)$ ,  $p(z)$ ,  $p(z/x)$  this key step may be easy or difficult to carry out. Several classes of distribution which have nice properties for this purpose can be found in<sup>1</sup>. The density function  $p(x/z)$  is known as the aposteriori density function of  $x$ . It is the knowledge about the state of nature after the measurements  $z$ . By definition, it contains all the information necessary for estimation.

- (iv) Depending on the criterion function for estimation one can compute estimate  $\hat{x}$  from  $p(x/z)$ . Some typical examples are:

- a) Criterion: Maximize the Probability ( $\hat{x} = x$ )

$$\text{Solution : } \hat{x} = \text{Mode of } p(x/z) \quad (5)$$

This is defined as the Most Probable Estimate

When the apriori density function  $p(x)$  is uniform, this estimate is identical to the classical maximum likelihood estimate.

b) Criterion: Minimize  $\int ||x - \hat{x}||^2 p(x/z) dx$

$$\text{Solution : } \hat{x} = E(x/z)^* \quad (6)$$

This is the conditional mean estimate.

c) Criterion: Minimize Maximum  $|x - \hat{x}|$

$$\text{Solution : } \hat{x} = \text{Medium of } p(x/z) \quad (7)$$

This can be defined as the minimax estimate.

Pictorially, the three estimates are shown in Fig. 1 for a general  $p(x/z)$  for a scalar case

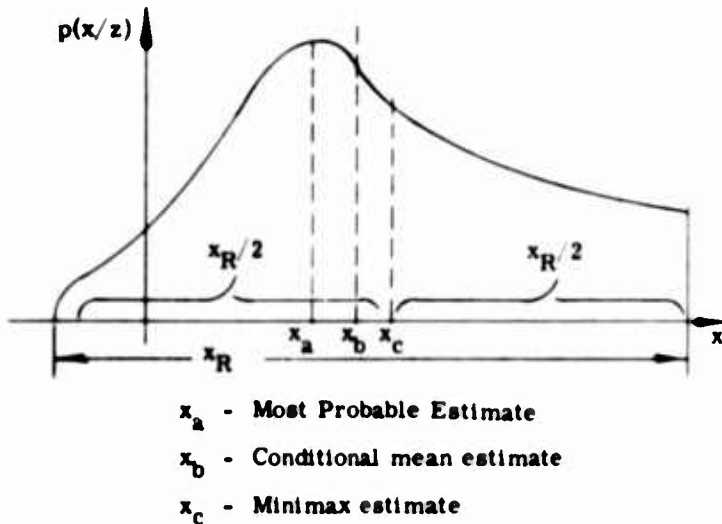


Fig. 1 Estimates based on aposteriori density

Clearly, other estimates, as well as confidence intervals can be derived from  $p(x/z)$  directly.

#### Special Case of the Wiener-Kalman Filter (single stage)

Now a special case of the above estimation problem will be considered. Let there be given

(i) A set of measurements  $z = (z_1, z_2, \dots, z_k)$

(ii) The physical relationship

$$z = Hx + v \quad (8)$$

(iii) The independent noise and state density functions

$$p(x, v) = p(x) p(v) \quad (9)$$

$$\left. \begin{array}{l} p(x) \text{ be gaussian with} \\ E(x) = \bar{x} \\ \text{Cov}(x) = P_0 \end{array} \right\} \quad (10)$$

$$\left. \begin{array}{l} p(v) \text{ be gaussian with} \\ E(v) = 0 \\ \text{Cov}(v) = R \end{array} \right\} \quad (11)$$

Now following the steps for the Bayesian solution one has

(i) Evaluate  $p(z)$ .

Since  $z = Hx + v$  and  $x, v$  is gaussian and independent, one immediately gets

$$\left. \begin{array}{l} p(z) \text{ is gaussian} \\ E(z) = H\bar{x} \\ \text{Cov}(z) = H P_0 H^T + R \end{array} \right\} \quad (12)$$

(ii) Evaluate  $p(x, z)$ .

Since  $\left(\frac{\partial}{\partial z}\right) = \text{Identify matrix}$ , it follows

$$\begin{aligned} p(x, z) &= p(x, v = z - Hx) \\ &= p(x) p_v(z - Hx)^{**} \end{aligned} \quad (13)$$

(iib) Evaluate  $p(z/x)^*$

$$p(z/x) = \frac{p(x, z)}{p(x)} = p(v) = p_v(z - Hx) \quad (14)$$

(iii) Evaluate  $p(x/z)$ .

One gets from Bayes' rule,

$$p(x/z) = \frac{p(x) p(v)}{p(z)} \quad (15)$$

By direct substitution of (10), (11), and (12), one obtains

$$\begin{aligned} p(x/z) &= \frac{|H P_0 H^T + R|^{\frac{1}{2}}}{(2\pi)^{n/2} |P_0|^{\frac{1}{2}} |R|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} [(x - \bar{x})^T P_0^{-1} (x - \bar{x}) \right. \\ &\quad + (z - Hx)^T R^{-1} (z - Hx) \\ &\quad \left. - (z - Hx)^T (H P_0 H^T + R)^{-1} (z - Hx)] \right\} \end{aligned} \quad (16)$$

Now completing squares in the  $\{ \}$ , (16) simplifies to

$$p(x/z) = \frac{|H P_0 H^T + R|^{\frac{1}{2}}}{(2\pi)^{n/2} |P_0|^{\frac{1}{2}} |R|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x - \hat{x})^T P^{-1} (x - \hat{x}) \right\} \quad (17)$$

where

$$P^{-1} = P_0^{-1} + H^T R^{-1} H \quad (18)$$

or equivalently,

$$P = P_0 - P_0 H^T (H P_0 H^T + R)^{-1} H P_0 \quad (19)$$

and

$$\hat{x} = \bar{x} + P H^T R^{-1} (z - H\bar{x}) \quad (20)$$

(iv) Now since  $p(x/z)$  is gaussian, the most probable, conditional mean, and minimax estimate all coincide and is given by  $\hat{x}$ .

This is the derivation of the single stage Wiener-Kalman filter<sup>2,3</sup>. The pair  $(P, \hat{x})$  is called a sufficient statistic for the problem in the sense that  $p(x/z) = p(x/P, \hat{x})$ .

#### MULTI-STAGE ESTIMATION PROBLEM

The problem formulation and the solution in this case is basically similar to the single stage problem. The only additional complication is that now the state is changing from stage to stage according to some dynamic relationship and that the aposteriori density function is to be computed recursively.

#### Problem Statement

It is assumed that at any stage  $k + 1$ , the following data is given as a result of previous computation or as part of the problem statement:

\*Note: step (iib) is redundant.

\*\* $p_v(z - Hx)$  means substituting  $(z - Hx)$  for  $v$  in  $p(v)$ .

\*It is assumed that  $p(x/z)$  has finite second moment.

- (i) The system equations governing the evaluation of the state.

$$\begin{aligned} x_{k+1} &= f(x_k, w_k) \\ z_{k+1} &= h(x_{k+1}, v_{k+1}) \end{aligned} \quad (21)$$

where  $x_{k+1}$  is the state vector at  $k+1$ .

$v_{k+1}$  is the measurement noise at  $k+1$ .

$z_{k+1}$  is the additional measurement available at  $k+1$ .

$w_k$  is the disturbance vector at  $k$ .

- (ii) The complete set of measurements  $Z_{k+1} \triangleq (z_1, \dots, z_{k+1})$ .

- (iii) The density functions\*

$$p(x_k/z_1, \dots, z_k) \triangleq p(x_k/Z_k)$$

$p(w_k, v_{k+1}/x_k)$  - statistics of a vector random sequence with components  $w_k$  and  $v_{k+1}$  which depends on  $x_k$ .

Now it is required to estimate  $x_{k+1}$  based on measurements  $z_1, \dots, z_{k+1}$ .

#### The Bayesian Solution

The procedure is analogous to the single stage case.

- (i) Evaluate  $p(x_{k+1}/x_k)$ . This can be accomplished either experimentally or analytically from knowledge of  $p(w_k, v_{k+1}/x_k)$ ,  $p(x_k/Z_k)$  and (21).

- (ii) Evaluate  $p(z_{k+1}/x_k, x_{k+1})$ . This is derived from  $p(w_k, v_{k+1}/x_k)$  and (21).

- (iii) Evaluate

$$\begin{aligned} p(x_{k+1}, z_{k+1}/Z_k) &= \int p(z_{k+1}/Z_k, x_{k+1}) \\ &\quad p(x_{k+1}/x_k) \cdot p(x_k/Z_k) dx_k \end{aligned} \quad (22)$$

From this the marginal density functions  $p(x_{k+1}/Z_k)$  and  $p(z_{k+1}/Z_k)$  can be directly evaluated.

- (iv) Evaluate

$$p(x_{k+1}/Z_{k+1}) = \frac{p(x_{k+1}, z_{k+1}/Z_k)}{p(z_{k+1}/Z_k)} \quad (23)$$

from (22)

$$= \frac{\int p(z_{k+1}/Z_k, x_{k+1}) p(x_{k+1}/x_k) p(x_k/Z_k) dx_k}{\int \int p(z_{k+1}/Z_k, x_{k+1}) p(x_{k+1}/x_k) p(x_k/Z_k) dx_{k+1} dx_k} \quad (24)$$

Eqn. (24) is a functional integral difference equation governing the evaluation of the a posteriori density function of the state of (21).

- (v) Estimates for  $x_{k+1}$  can now be obtained from  $p(x_{k+1}/Z_{k+1})$  exactly as in the single stage case.

#### Special Case of the Wiener-Kalman Filter\*

The given data at  $k+1$  is specified as follows:

The physical model is given by

$$\begin{aligned} x_{k+1} &= \Phi x_k + \Gamma w_k \\ z_k &= H x_k + v_k \end{aligned} \quad (25)$$

where  $w$  and  $v$  are independent, white, gaussian random sequences with

$$\begin{aligned} p(x_k/Z_k) &\text{ is Gaussian} \\ E(x_k/Z_k) &\triangleq \hat{x}_k \\ \text{Cov}(x_k/Z_k) &= P_k \end{aligned} \quad (26)$$

$$\begin{aligned} p(w_k, v_{k+1}/x_k, Z_k) &= p(w_k, v_{k+1}) = p(w_k) p(v_{k+1}) \\ E(v_k) &= E(v_{k+1}) = 0 \\ \text{Cov}(w_k) &= Q; \text{Cov}(v_{k+1}) = R \end{aligned} \quad (27)$$

Since in this case, the noise  $w_k, v_{k+1}$  is not dependent on the state, Eqn. (24) simplifies to

$$p(x_{k+1}/Z_{k+1}) = \frac{p(z_{k+1}/x_{k+1})}{p(z_{k+1}/Z_k)} p(x_{k+1}/Z_k) \quad (24)'$$

Hence the solution only involved the evaluation of the three density functions on the r.h.s. of (24)' given the data (25 - 27). This is carried out below:

From (27), it is noted that  $p(x_{k+1}/Z_k)$  is gaussian and independent of  $v_{k+1}$

$$\begin{aligned} E(x_{k+1}/Z_k) &= \Phi \hat{x}_k \\ \text{Cov}(x_{k+1}/Z_k) &= \Phi P_k \Phi^T + \Gamma Q \Gamma^T \triangleq M_{k+1} \end{aligned} \quad (28)$$

Similarly,  $p(z_{k+1}/Z_k)$  is gaussian and

$$\begin{aligned} E(z_{k+1}/Z_k) &= H \Phi \hat{x}_k \\ \text{Cov}(z_{k+1}/Z_k) &= H M_{k+1} H^T + R \end{aligned} \quad (29)$$

Finally  $p(z_{k+1}/x_{k+1})$  is also gaussian with

$$\begin{aligned} E(z_{k+1}/x_{k+1}) &= H x_{k+1} \\ \text{Cov}(z_{k+1}/x_{k+1}) &= R \end{aligned} \quad (30)$$

\*Footnote added in proof.

This development of the multistage Wiener-Kalman filtering method is very similar to a paper by Drs. H. Rauch, F. Tung, and C. Striebel entitled "On The Maximum Likelihood Estimate for Linear Dynamic Systems" presented at the SIAM Conference on System Optimization, 1964, Monterey, California. The only difference between the two developments is this: The Rauch-Tung-Striebel paper does not explicitly compute  $p(x/z)$  but simply computes its maximum and uses it as the estimate. In the author's approach, the computation of the maximum plays a secondary role. The explicit calculation of the posteriori probability is emphasized as the Bayesian viewpoint.

The authors are indebted to Prof. A. E. Bryson for bringing this reference to their attention.

\*The product of the two density functions yields  $p(w_k, v_{k+1}, x_k/Z_k)$  by the markov property of (21). It is also assumed that if  $p(w, v/x) = p(w, v)$  then  $w, v$  is a white random sequence.

Combining (28-30) using (24)', one gets

$$p(x_{k+1}/Z_{k+1}) = \frac{|H M_{k+1} H^T + R|^{-1/2}}{(2\pi)^{n/2} |R|^{-1/2} |M_{k+1}|^{-1/2}} \exp \left\{ -\frac{1}{2} [(x_{k+1} - \hat{x}_k)^T M_{k+1}^{-1} (x_{k+1} - \hat{x}_k) + (z_{k+1} - H \hat{x}_{k+1})^T R^{-1} (z_{k+1} - H \hat{x}_{k+1}) - (z_{k+1} - H \hat{x}_k)^T (H M_{k+1} H^T + R)^{-1} (z_{k+1} - H \hat{x}_k)] \right\} \quad (31)$$

Now completing squares in  $\{ \}$  one gets,

$$p(x_{k+1}/Z_{k+1}) = \frac{|H M_{k+1} H^T + R|^{-1/2}}{(2\pi)^{n/2} |R|^{-1/2} |M_{k+1}|^{-1/2}} \exp \left\{ -1/2 (x_{k+1} - \hat{x}_{k+1})^T P_{k+1}^{-1} (x_{k+1} - \hat{x}_{k+1}) \right\} \quad (32)$$

where

$$\hat{x}_{k+1} = \hat{x}_k + M_{k+1} H^T (H M_{k+1} H^T + R)^{-1} (z_{k+1} - H \hat{x}_k) \quad (33)$$

$$P_{k+1}^{-1} = M_{k+1}^{-1} + H^T R^{-1} H \quad (34)$$

or equivalently

$$P_{k+1} = M_{k+1} - M_{k+1} H^T (H M_{k+1} H^T + R)^{-1} H M_{k+1} \quad (35)$$

and

$$M_{k+1} = P_k + Q \quad (36)$$

Eqns. (33-36) are exactly the discrete Wiener-Kalman filter in the multistage case [3] [4].

#### A SIMPLE NONLINEAR NONGAUSSIAN ESTIMATION PROBLEM

The discussions in the above sections have been carried out in terms of continuous density functions. However, it is obvious that the same process can be applied to problems involving discrete density function and discontinuous functional relationships. It is worthwhile, at this point, to carry out one such solution for a simple contrived example which nevertheless illustrates the application of the basic approach.

The problem can be visualized as an abstraction of the following physical estimation problem. An infrared detector followed by a threshold device is used in a satellite to detect hot targets on the ground. However, extraneous signals, particularly reflection from clouds, obscure the measurements. The problem is to design a multistage estimation process to estimate the presence of hot targets on the ground through measurement of the output of the threshold detector.

Let  $s_k$  (target) be a scalar independent Bernoulli process with,

$$p(s_k) = (1-q) \delta(s_k) + q \delta(1-s_k) \quad (37)^*$$

\*The notation  $\delta(x) = \begin{cases} 1 & x=0 \\ 0 & x \neq 0 \end{cases}$  is used here. Also,  $p(x)$  is to be interpreted as mass functions.

$n_k$  (cloud noise) be a scalar Markov process with,

$$p(n_1) = (1-a) \delta(n_1) + a \delta(1-n_1) \quad (38)$$

$$p(n_{k+1}/n_k) = (1-a - \frac{n_k}{2}) \delta(n_{k+1}) + (a + \frac{n_k}{2}) \delta(1-n_{k+1}) \quad (39)$$

and the scalar measurement,

$$z_k = s_k \odot n_k \quad (40)$$

where  $\odot$  indicates the logical "OR" operation.

Essentially Eqns. (37-40) indicate the fact that as the detector sweeps across the field of view, cloud reflection tends to appear in groups while targets appear in isolated dots.

Now we proceed to the Bayesian solution. First, we have,

$n_1$	0	0	1	1
$s_1$	0	1	0	1
$z_1$	0	1	1	1
Probability of $z_1$	$(1-a)(1-q)$	$q(1-a)$	$a(1-q)$	$aq$

$$p(z_1) = (1-a)(1-q) \delta(z_1) + (a+q-aq) \delta(z_1-1) \quad (41)$$

Also,

$$p(z_1/n_1) = \delta(z_1-1) n_1 + [(1-q) \delta(z_1) + q \delta(z_1-1)] (1-n_1) \quad (42)$$

Then by direct calculation,

$$p(n_1/z_1) = \frac{p(z_1/n_1) p(n_1)}{p(z_1)} = (1-a'(z_1)) \delta(n_1) + a'(z_1) \delta(n_1-1)$$

where

$$a'(z_1) = \frac{a \delta(z_1-1)}{(1-a)(1-q) \delta(z_1) + (a+q-aq) \delta(z_1-1)} \quad (44)$$

Similarly,

$$p(z_1/s_1) = \delta(z_1-1) s_1 + [(1-a) \delta(z_1) + a \delta(z_1-1)] (1-s_1) \quad (45)$$

and

$$p(s_1/z_1) = \frac{p(z_1/s_1) p(s_1)}{p(z_1)} = (1-q'(z_1)) \delta(s_1) + q'(z_1) \delta(s_1-1) \quad (46)$$

where

$$q'(z_1) = \frac{q \delta(z_1-1)}{(1-a)(1-q) \delta(z_1) + (a+q-aq) \delta(z_1-1)} \quad (47)$$

and a reasonable estimate is

$$\hat{s}_1 = \begin{cases} 1 & \text{if } q'(z_1) \geq \epsilon \text{ (Given constant)} \\ 0 & \text{if } q'(z_1) < \epsilon \end{cases} \quad (48)$$

where  $\hat{s}_1 = 1$  may be interpreted as an alarm.

Now consider a second measurement  $z_2$  has been made. One has,

$$p(n_2/z_1) = \int_{-\infty}^{\infty} p(n_2/n_1) p(n_1/z_1) dn_1 \quad (49)$$

which after straightforward but somewhat laborious manipulations becomes,

$$= (1-a - \frac{a'(z_1)}{2}) \delta(n_2) + (\frac{a'(z_1)}{2} + a) \delta(n_2 - 1)$$

$$\triangleq (1-a(z_1)) \delta(n_2) + a(z_1) \delta(n_2 - 1)$$

Furthermore,

$$p(s_2/z_1) \triangleq p(s_2) = (1-q) \delta(s_2) + q \delta(s_2 - 1) \quad (50)$$

Eqns. (49) and (50) now take the place of (37) and (38) and by the same process, one can get in general,

$$p(n_k/Z_k) \triangleq p(n_k/z_1, z_{k-1}, \dots) \\ = (1-a'(Z_k)) \delta(n_k) + a'(Z_k) \delta(n_k - 1) \quad (51)$$

$$a'(Z_k) \triangleq a'(z_k, z_{k-1}, \dots)$$

$$= \frac{a(z_{k-1}) \delta(z_{k-1})}{(1-a(z_{k-1}))(1-q) \delta(z_k) + (a(z_{k-1}) + q - a(z_{k-1})q) \delta(z_k - 1)} \quad (52)$$

$$a(Z_{k-1}) \triangleq a(z_{k-1}, z_{k-2}, \dots) \\ = a + \frac{a'(Z_{k-1})}{2} \quad (53)$$

$$p(s_k/Z_k) \triangleq p(s_k/z_k, z_{k-1}, \dots) \\ = (1-q'(Z_k)) \delta(s_k) + q'(Z_k) \delta(s_k - 1) \quad (54)$$

$$q'(Z_k) \triangleq q'(z_k, z_{k-1}, \dots) \\ = \frac{q \delta(z_k - 1)}{(1-a(z_k))(1-q) \delta(z_k) + (a(z_{k-1}) + 1 - a(z_{k-1})q) \delta(z_k - 1)} \quad (55)$$

$$p(n_{k+1}/Z_k) = (1-a(Z_k)) \delta(n_{k+1}) + a(Z_k) \delta(n_{k+1} - 1) \quad (56)$$

$$p(s_{k+1}/Z_k) = p(s_{k+1}) \quad (57)$$

Eqns. (51-57) now represent the general recursion solution for the multistage estimation process.

As a check, two possible observed sequences for  $z$ , namely (0, 1) and (1, 1) are considered. With  $a = 1/4$  and  $q = 1/4$  it is found that  $p(s_2/z_2, z_1) = 0.571$  and  $0.337$  respectively. This agrees with intuition since the sequence (1, 1) has a higher probability of being cloud reflections. On the other hand, the numbers also showed that under the circumstances, it is very difficult to detect targets with accuracy using the system contrived here.

Often times one is actually interested in  $p(s_k/Z_{k,\tau})$  with  $\tau > 0$  in order to obtain the so-called "smoothed" estimate for  $s_k$ . The desired density function can be computed from  $p(s_k/Z_k)$  by further manipulations. However, the calculation becomes involved and will not be done here.

## RELATIONSHIP TO GENERAL BAYESIAN STATISTICAL DECISION THEORY

It is worthwhile to point out the relationship of the above formulation and solution of the estimation problem to and its difference from the general statistical decision problem. For simplicity, the single stage case is considered again. In the general statistical decision problem, the input data is somewhat different. One typical form is:

$p(x)$  - a priori density of  $x$

$\{e\}$  - a set of choices of experiments from which we can derive measurements  $z$  with

$p(z/x, e)$ -conditional density of  $z$  for given  $x$  and  $e$ .

$\{u\}$  - a set of choices of decisions

$J(e, z, u, x)$ - a criterion function which is a possible function of  $e, z, u$  and  $x$ .

The problem is then stated as the determination of  $e$  and  $u$  so that  $E(J)$  is optimized. The optimal  $J$  is given by

$$J_{\text{opt}} = \text{Max}_e (\text{Min}_u) \int \left\{ \text{Max}_u (\text{Min}_e) \left[ \int J(e, z, u, x) \cdot p(x/z, e) dx \right] p(z/e) dz \right\} \quad (58)**$$

Thus, the main differences between the estimation problem and the general decision problem are as follows:

(i) In the estimation problem there is no choice of experiment. One always makes the same type of measurement  $z$  given by  $g(x, v)$ . To generalize the estimation problem, one can specify:

$$z_e = g_e(x, v); \{e\} = 1, 2, \dots = \text{possible sets of measurements} \quad (59)$$

and then require that

$$\hat{x} = \text{Opt}_e \{(\hat{x})_e \mid e = 1, 2, \dots\}$$

(ii) In the general decision problem, the function  $z = g(x, v)$  is implicit in  $p(z/x, e)$ . Hence step (iia) and (iib) for the estimation solution is not required. This often is a tremendous simplification.

(iii) In the estimation problem the criterion function  $J$  is always a simple function of  $x$  only. There is, furthermore, no choice of action (one has to make an estimate by definition). On the other hand, the general decision problem is more analogous to a combined estimation and control problem where one has a further choice of action after determining  $p(x/z)$ , and like a control problem, the criteria function is generally more complex.

(iv) It is, however, to be noted that the key step is the computation of  $p(x/z)$  for both problems. The choice of action is determined only after the computation of  $p(x/z)$ . Thus, a general decision problem can be decomposed into two problems, namely, determination of  $p(x/z)$  (estimation problem) and choice of action (control problem). In control-theoretic technology, this fact is called the Generalized Decomposition Axiom.

As an example, consider the single stage Wiener-Kalman problem. The added requirement that,

$$J(e, z, u, x) = J(u, x) = E \| Bx + u \|^2 \\ = \int \| Bx + u \|^2 p(x/z) dx \quad (60)$$

\*For other equivalent forms see [1]

\*\*See [1]

be a minimum. Expanding (60), one gets

$$J = E \| x \|^2_{B^T B} + 2 u^T B \hat{x} + u^T u \quad (61)$$

Clearly,

$$u_{\text{opt}} = u(\hat{x}) \triangleq u(x(z)) \triangleq u(z) = -B \hat{x}(z) \quad (62)$$

which is one of the fundamental results of linear stochastic control. Thus, the control action  $u$  is only a function of the criterion  $J$  and the a posteriori density function  $p(x/z)$ . In fact, in this case only  $\hat{x}$  of  $p(x/z)$  is needed. We call  $\hat{x}$  as the minimal sufficient statistic for the control problem.

In the more general multistage case, the decomposition property clearly still holds. The only difference being that  $p(x_{k+1}/Z_{k+1})$  is now dependent on  $u_k$ . However, this dependence is entirely deterministic since in a given situation one always knows what  $u_k$ 's are. In fact, in the Wiener-Kalman control problem, it is known that  $u_k$  is a linear function of  $\hat{x}_k$  only.

### CONCLUSION

In the above sections, the problem of estimation from the Bayesian viewpoint is discussed. It is the author's thesis that this approach offers a unifying methodology, at least conceptually, to the general problems of estimation and control.

The a posteriori conditional density function  $p(x/z)$  is seen to be the key to the solution of the general problem. Difficulties associated with the solution of the general problem now appears more specifically as difficulties in steps leading to the computation of  $p(x/z)$ . From the above discussions, it is relatively obvious that these difficulties are:

(i) Computation of  $p(z/x)$  -

In both the single stage or multistage case, this problem is complicated by the nonlinear functional relationships between  $z$  and  $x$ . Except in the case when  $z$  and  $x$  are linearly related or when  $z$  and  $x$  are scalars, very little can be done in general, analytically or experimentally. As was mentioned earlier, this difficulty does not appear in the usual decision problem, since there it is assumed that  $p(z/x)$  is given as part of the problem.

(ii) Requirement that  $p(x/z)$  be in analytical form.

This is an obvious requirement if we intend to use the solution in real-time applications. It will not be feasible to compute  $p(x/z)$  after  $z$  has occurred.

(iii) Requirement that  $p(x)$ ,  $p(z)$ ,  $p(x/z)$  be conjugate distributions.\*

This is simply the requirement that  $p(x)$  and  $p(x/z)$  be density functions from the same family. Note that all the examples discussed in this paper possess this desirable property. This is precisely the reason that multistage computation can be done efficiently. This imposed a further restriction on the functions  $g$ ,  $f$  and  $h$ .

The difficulties (i - iii) listed above are formidable ones. It is not likely that they can be easily circumvented except for special classes of problems, such as those discussed. However, it is worthwhile first to pinpoint these difficulties. Researches toward their solution can then be effectively initiated. Finally, it is felt that the Bayesian approach offers a unified and intuitive viewpoint particularly adaptable to handle modern day control problems where the state and the Markov assumptions play a fundamental role.

### REFERENCES

- [1] H. Raiffa, and R. Schlaifer, "Applied Statistical Decision Theory", Harvard University Press, 1961, Ch. III.
- [2] A. E. Bryson and M. Frazier, "Smoothing in Linear and Nonlinear Systems", Proc. of Optimal System Synthesis Symposium, Sept. 1962, Wright Field, Ohio ASD-TDR-63-119.
- [3] Y. C. Ho, "On Stochastic Approximation and Optimal Filtering Method", Journal of Mathematical Analysis and Applications, February 1963.
- [4] R. E. Kalman, "New Methods and Results in Linear Prediction and Filtering Theory", RIAS Report 61-1.

\*See [1]



## BASIC TECHNICAL REPORT LIST\*

Defense Documentation Center  
Cameron Station [20]  
Alexandria, Virginia 22314  
Attn: DDC-IRA

Director  
U. S. Army Electronics Research  
and Development Laboratories  
Fort Monmouth, New Jersey [24]  
Attn: AMSEL-RD-DR AMSEL-RD-

K	SR
XE	SS
XC	P
XS	PE
N	PR
NR	PF
NE	G
ND	GF
NP	GD
S	ADT
SA	FU 1
SE	

U. S. Army Research Office  
3045 Columbia Pike  
Arlington, Virginia  
Attn: Physical Sciences Division

U. S. Army Research Office (Durham)  
P. O. Box CM, Duke Station  
Durham, North Carolina  
Attn: Dr. H. Rohl

Commanding General  
U. S. Army Materiel Command  
Washington 25, D. C.  
Attn: AMCRD-RS-PE-E

Commanding Officer  
Harry Diamond Laboratories  
Connecticut Ave. and Van Ness St. N. W.  
Washington, D. C.

Commanding General  
U. S. Army Electronics Research  
and Development Activity  
Fort Huachuca, Arizona 85611

Commanding General  
U. S. Army Engineers Research and  
Development Laboratory  
Fort Belvoir, Virginia 22060

Commanding General  
Watertown Arsenal  
Watertown, Massachusetts  
Attn: U. S. Army Material Research  
Agency

Commanding General  
U. S. Army Ballistics Research Lab.  
Aberdeen Proving Ground  
Aberdeen, Maryland 21005

Commanding Officer  
Limited Warfare Laboratory  
Aberdeen Proving Ground  
Aberdeen, Maryland 21005  
Attn: Technical Director

Commanding Officer  
U. S. Army Missile Command  
Redstone Arsenal, Alabama 35809

Commanding Officer  
U. S. Army Electronics Research and  
Development Activity  
White Sands, New Mexico

Commanding Officer [10]  
Office of Naval Research  
Navy 100, Box 39  
Fleet Post Office  
New York, New York

Commanding General  
Frankford Arsenal  
Philadelphia, Pennsylvania  
Attn: Dr. Sidney Ross

Director  
Advanced Research Projects Agency  
Washington 25, D. C.

Commanding Officer  
U. S. Army Security Agency  
Washington 25, D. C.

Chief of Naval Research [2]  
Department of the Navy  
Washington, D. C. 20360  
Attn: Code 427 Dr. Shostak

Commanding Officer [2]  
Office of Naval Research  
495 Summer Street  
Boston, Massachusetts 02110

Chief of Naval Research  
Department of the Navy  
Washington, D. C. 20360  
Attn: Code 421

Commanding Officer  
Office of Naval Research  
230 N. Michigan Avenue  
Chicago, Illinois 60601

Commanding Officer  
Office of Naval Research  
207 West 24th Street  
New York, New York

Commanding Officer  
Office of Naval Research  
1030 East Green Street  
Pasadena, California

Commanding Officer  
Office of Naval Research Branch Office  
1000 Geary Street  
San Francisco, California 94109

The Director [6]  
Naval Research Laboratory  
Washington 25, D. C.  
Attn: Technical Information Office

Assistant Secretary of Defense  
(Research and Development)  
Research and Development Board  
Department of Defense  
Washington 25, D. C.

Director  
U. S. Naval Research Laboratory  
Washington, D. C. 20390  
Attn: Code 2027

Martin A. Garstens  
Magnetism Branch, Code 6450  
Solid State Division  
Naval Research Laboratory  
Washington 25, D. C.

Chief of Naval Operations  
Department of the Navy  
Washington 25, D. C.  
Attn: Op-20

Chief of Naval Operations  
Department of the Navy  
Washington 25, D. C.  
Attn: Op-32

Chief of Naval Operations  
Department of the Navy  
Washington, D. C. 20380  
Attn: Op-07T

Hq. AFRL (CRXL) [5]  
L. G. Manscom Field  
Bedford, Massachusetts 01731  
[1] CRXL  
[2] CRR-Dr. Hollingsworth  
[3] CRF  
[4] CRW

AMRL (Library) 4570  
Wright-Patterson AFB, Ohio

Chief, Bureau of Ships [2]  
Department of the Navy  
Washington 25, D. C.  
Attn: Code 810

Commander [2]  
U. S. Naval Electronics Laboratory  
San Diego, California

Director  
Naval Ordnance Laboratory  
White Oak, Maryland

Technical Library  
U. S. Naval Proving Ground  
Dahlgren, Virginia

Commanding Officer  
U. S. N. Air Development Center  
Johnsville, Pennsylvania 18974  
Attn: NADC Library

Librarian  
U. S. Naval Post Graduate School  
Monterey, California 93940  
Attn: Technical Reports Section

Commanding General [3]  
Air Research and Development Command  
P. O. Box 1395  
Baltimore 3, Maryland  
Attn: RDTRRP

Director, Air University [2]  
Library  
Maxwell Air Force Base  
Alabama

RADC (EMLAL-1)  
Griffies Air Force Base  
New York 13442  
Attn: Documents Library

Air Force Office of Scientific Research  
Air Research and Development Command  
Washington 25, D. C.  
Attn: SRV, Physics Division

Commandant  
Air Force Institution of Technology  
Wright-Patterson Air Force Base  
Ohio  
Attn: Library

Systems Engineering Group  
Deputy for Systems Engineering  
Directorate of Technical Publications  
and Specifications (SEPRR)  
Wright-Patterson Air Force Base,  
Ohio 45433

AF Special Weapons Center  
Kirtland Air Force Base  
Albuquerque, New Mexico  
Attn: SWOI

AFMTC Technical Library  
Patrick AFB, Florida 32925

European Office, OAR  
Technical Information Office  
APO 667, New York  
New York

Headquarters  
Air Research and Development Command  
United States Air Force  
Andrews Air Force Base  
Washington 25, D. C.

Mr. A. D. Bedrosian  
Signal Corps Liaison Office  
Massachusetts Institute of Technology  
Building 26, Room 131  
Cambridge, Massachusetts 02139

Commanding Officer  
U. S. Army Research Office  
Box CM, Duke Station  
Durham, North Carolina  
Attn: CRS-AA-IP, Mr. Ulich

Commanding Officer  
U. S. Army Signal Missile  
Support Agency  
White Sands Missile Range  
New Mexico  
Attn: SICWS-MEW  
Mr. T. S. Bellows

Commanding Officer  
U. S. Army Engineer  
Research and Development Laboratories  
Fort Belvoir, Virginia 22060  
Attn: Technical Document Center

National Security Agency  
Physical Sciences Division  
Fort George Meade, Maryland  
Attn: Dr. Alvin Mackler

Dr. H. Compaigne  
National Security Agency  
Physical Sciences Division  
Fort George Meade, Maryland

Sandia Corporation  
Org. 1424 Sandia Base  
Albuquerque, New Mexico  
Attn: Dr. C. W. Harrison, Jr.  
Sandia Corporation  
Sandia Base  
Albuquerque, New Mexico  
Attn: Library Division 1922-1

Librarian  
National Bureau of Standards Library  
Room 301, Northwest Building  
Washington, D. C. 20234

Librarian  
U. S. Department of Commerce  
National Bureau of Standards  
Boulder, Colorado 80301

Massachusetts Institute of Technology  
Engineering Library  
Building 10 - Room 950  
Cambridge, Massachusetts 02139  
Attn: E. Farkas

U. S. Coast Guard  
1300 E. Street N. W.  
Washington 25, D. C.  
Attn: EEE

Stanford Electronics Laboratory  
Stanford University  
Stanford, California 94305  
Attn: Document Library  
Applied Electronics Laboratory

Lincoln Laboratory  
Massachusetts Institute of Technology  
P. O. Box 71  
Lexington, Massachusetts 02173  
Attn: Library A-082

Mr. John Hewitt  
Document Office, Room 26-327  
Research Laboratory of Electronics  
Massachusetts Institute of Technology  
Cambridge, Massachusetts 02139

Engineering Division  
Case Institute of Technology  
Cleveland, Ohio 44106  
Attn: H. R. Nara  
Associate Director

Advisory Group of Electron Devices  
346 Broadway, 8th Floor  
New York, New York 10013

Antenna Laboratory  
Department of Electrical Engineering  
The Ohio State University  
2024 Neil Avenue  
Columbus, Ohio 43210  
Attn: Reports Librarian

Polytechnic Institute of Brooklyn  
Graduate Center Library  
Route 110  
Farmingdale, L. I., New York 11735

The University of New Mexico  
Department of Electrical Engineering  
Albuquerque, New Mexico 87 56  
Attn: A. R. Koschmann, Chairman

Carlyle Barton Laboratories  
Johns Hopkins University  
Charles and 34th Street  
Baltimore, Maryland 21218  
Attn: Librarian

Battelle Memorial Institute  
505 King Avenue  
Columbus, Ohio 43201  
Attn: G. J. Falkenbach

Charles C. H. Tang  
Bell Telephone Laboratories 2D-347  
Murray Hill, New Jersey 07971

Prof. A. W. Lawson  
Department of Physics  
University of California  
Riverside, California

Gordon McKay Library [2]  
Pierce Hall  
Harvard University  
Cambridge, Massachusetts 02138

Chung Kung University  
Department of Electrical Engineering  
Tainan, Taiwan  
Republic of China  
Attn: Prof. Chao-Hai-Chou  
Head, English Department

for Contract Nonr-1866(28)

Department of Electrical Engineering  
King's College  
Newcastle upon Tyne  
England

for Contract Nonr-1866(32)

U. S. Atomic Energy Commission  
Office of Technical Information Extension  
P. O. Box 62  
Oak Ridge, Tennessee

\* One copy to each address unless otherwise  
indicated by numbers in brackets.